International Journal of Computer Science Engineering
and Information Technology Research (IJCSEITR)
ISSN (P): 2249–6831; ISSN (E): 2249–7943
Vol. 12, Issue 1, Jun 2022, 83–88

TRANS STELLAR
•Journal Publications •Research Consultancy

BLOCKCHAIN TECHNOLOGY: A REVOLUTIONARY SOLUTION OF BIG DATA CHALLENGES

MS. DEEPIKA KIRTI & DR. RUCHI SAWHNEY

Bosco Technical Training Society & Don Bosco Technical School, Delhi, India

ABSTRACT

In response to the growing acceptance of Blockchain technology, as well as the huge scope of its applications, a large amount of continuing research is being conducted in a number of practical and scientific disciplines in recent years. A revolutionary solution for modern technology concerns such as decentralization, trust, identification, data ownership, and data-driven decision-making, the Blockchain is being heralded despite the fact that it is still in its infancy and is still in the testing phase. The world is being forced to confront a digital data explosion as a result of a simultaneous increase in both the quantity and diversity of digital data being generated by humans and robots at the same time. Meanwhile, experts are constantly looking for the most effective means of storing, organizing, and analyzing large amounts of data, and Blockchain technology can provide significant assistance. The proposed solutions for Big Data in the areas of decentralized administration of private data, digital property resolution, IoT connectivity, and public institution reform, among other things, will have a significant impact on the future of the technology. An innovative solution to some of the Big Data fields that have not yet been thoroughly investigated is offered in this study, which makes use of Blockchain technology to accomplish this goal.

KEYWORDS: Blockchain, Decentralized, Big Data, IoT Connectivity, Distributed.

Received: Dec 22, 2022; **Accepted**: Jan 12, 2022; **Published**: Feb 11, 2022; **Paper Id**: IJCSEITRJUN20229

Original Article

INTRODUCTION

Since global data traffic has risen at an alarming rate over the few years, "big data" has become more popular. Health care, retail, transportation and manufacturing are among the numerous vertical industries where the use of big data is predicted to save money by 2025, when it is expected to reach a market value of $229.4 billion. For this reason, several sectors of research and engineering are investigating big data, including computer vision, Internet-of-things (IoT), performance management, and smart cities. In addition to the actual structure, Liu et al., (2019) analysed huge datasets from three perspectives: attribute, comparison, and structure. When it comes to analysing massive volumes of data, the term "big data" is becoming increasingly popular (e.g., high speed, data acquisition, and statistics). The term "big data" refers to vast amounts of data that can be stored, handled, analysed, or even photographed using ordinary internet procedures. According to the Big Data Institute, this type of information necessitates the use of a variety of horizontal measurement techniques in order to be properly processed structurally.

CHALLENGES IN BIGDATA

But there are a number of issues that are addressed by big data strategies and applications, such as data privacy and data security, as well as computer infrastructure monitoring and data management, as well as data interpretation and real-time data processing. Security and privacy are two of the most pressing concerns when dealing with big

data, owing to the prevalence of personal information it typically contains, such as age, address, demographics, preferences, and financial information. Various solutions and tactics have been examined in order to maintain the confidentiality of data and confidential information. For example, Liu et al., (2020) applied the same theory and integration game to tackle the resource allocation problem for protecting mobile social networks with massive data. It was demonstrated by Hu et al. that an algorithm for detecting the safety of an intelligent grid system could be created using augmented learning (2014). There are many advantages to using Blockchain and ledger technology to assure data security and privacy in large-scale operations. Industrial IoT uses of blockchain have demonstrated its ability to provide high-quality data while also securing data exchange. A blockchain-based approach was created to collect data from mobile ad networks and to facilitate the acquisition of mobile node data in research investigations. Additionally, the blockchain has been linked to edge servers to increase data quality and manage the complicated duties required by IoT devices while ensuring security. For the first time, blockchain can revolutionise current big data systems by enabling new big data apps and services by giving more efficient security features and the power of managed services. In this study, we examine everything from strategies to potential and future trends in the blockchain with big data.

Immutability

As a result of the blockchain's immutable network, it is nearly impossible to destabilise. A digital ledger is duplicated at each location, which makes it different from traditional financial systems. Before it is entered into the book, the nodes confirm and authorise the action after it has been initiated. Thus, the security and visibility of the system are dependent on the compatibility of all significant nodes before any transaction can be completed. The public blockchain, which allows everyone to verify transactions but prevents data alteration, is especially vulnerable to corruption because of this.

Decentralization

Rather than a single authority, the network is maintained by a group of nodes. As a result of the split process, users now have web-based access to the blockchain and the ability to save their duplicate data using private keys.

Security

The blockchain can provide high levels of security due to its fragile and constant nature. Unauthorized attacks can be thwarted with the use of cryptography's fire-fighting techniques. In order to protect the identity of each piece of data, it is accelerated, while at the same time providing its own unique identification. To ensure the data is unalterable, the hashes of the previous and current blocks are included in each block. Because of the hashing, the frame is also rigid. As a result, the public key cannot be converted to a private key, and harming the network necessitates fixing all data saved in each location.

Consensus

The alignment of alignment algorithms, which determines which nodes in a network are actively participating, is critical to the performance of blockchain frameworks. Using this, the verification process can be sped up and matched to the voting process.

Accelerated Financial Settlement

Traditional banking systems take a long time to process transactions. Foreign and visiting personnel can be relocated more quickly thanks to this technology. As a result of smart contracts, contract accounts can be managed much more swiftly.

DATA SECURITY BY BLOCKCHAIN

Apps that use large amounts of data have become increasingly popular, but they have also been plagued by severe security issues. The process of gathering data is a critical one in the whole data processing lifecycle. Suspicious data sources and communication channels can expose data gathering to numerous malicious attacks and threats. A safe data gathering technique is therefore required for all data applications. There have been a number of studies to ensure the safety of data collection to date. A blockchain-based massive data collection approach for mobile crowdsensing (MCS), for example, was introduced. Mobile terminals (MT) and sensors have made MCS a success in the industrial internet (IIoT) environment because of their expanding popularity. The MCS framework is built using cloud servers and a series of MTs. There are a number of services published by MCS servers related to hearing and selecting MTs for usage locally. Data collecting is hindered by a lack of MT power supply, sensor equipment, and secure data transfer. A model created by Wu et al., (2018) integrates blockchain with deep reinforcement learning to address these difficulties (DRL). Data is collected in a secure and energy-efficient manner from many sources. Data gathering and a variety of sensors are provided through a distributed blockchain-based DRL system for each MT. MTs transmit data, but the Ethereum blockchain technology ensures the integrity of that data. Additionally, Ethereum uses a separate business to maintain the book safe and share it with affiliate MTs. Multiple attacks, device failures, eclipse attacks, and more are all addressed by the proposed solution.

TRANSMISSION/SHARING OF SECURE BIG DATA

Using a solid blockchain ecosystem, data may be transferred safely. Secure data transfer from data sources to analysis is also made possible, addressing security and privacy issues that have long dogged traditional methods of data transfer. The blockchain can ensure big data training and avoid data theft to facilitate connections in big data. This data can be gathered through several means, such as through the use of social media, data libraries, and other support technologies. A hash signature and a value are then added to the data before it is shared with data analytic services, ensuring great transparency and trustworthiness in data sharing for all parties involved in the data flow.

FILE SYSTEMS SECURITY BY BLOCKCHAIN

There are a few cloud-based options for storing and retrieving files from any location. Many individuals and organisations are leery of storing confidential information on a third-party server. The cloud provider still faces some security difficulties, even with pre-loading content encryption. Medical professionals rely on electronic information systems the most (Pouyanfar et al., 2018). Every day, medical images, records, diagnostic reports, and other data are created. Electronic medical records can have an impact on patient care, and this data is shared with other medical institutions. The privacy of the patient is jeopardised if the supplied medical data is improperly used. Alternatives must be found in order to limit access to medical records. Integrating a blockchain with an interplanetary file system can address these kinds of security concerns (IPFS). IPFS is a last-ditch effort to address the issue of file idiocy. A hash value is generated for each saved file, making it possible for the user to locate the file by utilising the hash address. Medical data is secured using an attribute-based mechanism before it is stored in the cloud. The ciphertext is tied to the policy, while the user's private key is related to their attributes. Ciphertext encryption can be removed by any user if their private key satisfies the encryption's access policy. The blockchain is also used to keep track of the storage and retrieval of data. In order to demonstrate the legitimacy of user authentication, the blockchain stores the hash value of medical storage data. A single point of failure is eliminated thanks to the blockchain's fixed structure.

TRAINING OF SECURE DATA

The rise of edge and cloud computing has increased the volume of data in numerous settings. The use of numerous machine learning (ML) and deep learning (DL) algorithms is essential to a successful data analysis strategy. For its efficiency and accuracy, the SVM has become a prominent method for machine learning (ML). Automobile social networking sites, social media providers, auto manufacturers, and automotive management groups all provide data for this study. The data properties of different data sources can be rather diverse. SVM separator training suffers from a lack of data attributes due to a wide range of sources. Thus, data must be provided between several businesses to merge the database with multiple attributes and train separator. Data privacy is threatened by the sharing of data from multiple businesses. Blockchain-based SVM training for vertically partitioned datasets from several data providers has been established. The blockchain consortium and homomorphic cryptosystem were utilised to create a secure training platform that did not require the involvement of a third-party trusted authority. During the training phase, contacts between the entities involved in the process are protected by a homomorphic cryptosystem and a blockchain consortium. The blockchain consortium contributes to the success of communication by providing a public and secure data exchange environment for users.

PRIVACY PRESERVATION IN PROCESSING BIG DATA

The need to protect individual privacy has grown in importance in light of the tremendous rise in the production of data (Alam, 2021). Data privacy has emerged as a serious issue in recent years as a result of the massive expansion in data collection. In the current era of big data, data is always available and updated, which helps to increased sales and innovation (Zhang et al., 2020). In major organisations and corporations, the data is used to improve customer service and make better decisions. Since then, data has become a valuable commodity. To keep track of city traffic and upkeep, as well as to verify that the air and water are safe to breathe, big data is frequently employed in cities with advanced infrastructure. ITS (intelligent transportation systems) of automobile navigation systems in smart cities are protected by a blockchain-based paradigm. A blockchain-based storage system is used for offline storage, where all user privacy is protected. A standard key is assigned to a group of vehicles, which encrypts sensitive information using the same key. Sharing information about speeding and blocking, location enabling and disabling, and other safety features is possible.

SMART CITY BY BLOCKCHAIN

To create smart cities, rapid urbanisation has necessitated the development of intelligent transportation systems as well as solutions for environmental, administrative, and energy management issues. IoT, big data, and energy-efficient internet technologies can all work together to provide the infrastructure that a smart city needs. Security, dependability, care, flexibility, and cost are all key considerations. All of the IoT device's needs are met by blockchain technology, including transparency, power efficiency, space use, rediscovery, and storage. Asymmetric encryption, algorithm compliance, blockchain structure, and the Merkle tree acquisition function for deconstruction are all described in Manyika's presentation (2011). With the use of the Merkle tree, the building is even more secure when it comes to accomplishing discreet duties because the blocks are connected to each other within the block itself Large-scale data-based auditing solutions known as external auditing firms have been developed in the last several years (TPAs). TPAs are vulnerable to cloud security since they are intermediary entities. Blockchain technology was used to develop TPA-designated smart cities that are more secure and reliable.

SMART HEALTHCARE

Healthcare sector advances have led to an increase in the amount of medical data being collected. A patient's diagnosis and treatment can be made more accurate with this information. Increasingly, healthcare experts are focusing on IoT and wearable technologies, where sensors, equipment and vehicles connected to the internet deliver services for the greater good of humanity. Elderly patients are routinely monitored by remote access, for example (Su et al., 2017). Data communication and monitoring are both safer with these new technologies. Because of this, data security and privacy could be jeopardised by these dangers. Adopting a blockchain as a solution provides data analysis security and efficiency, but it is costly and lags behind in terms of adoption.

SMART TRANSPORTATION

An efficient mode of transportation can convey people and things between locations. The transportation business might greatly benefit from blockchain deployment, yet many in the industry are unfamiliar with this new technology. MaaS, Internet of Things (IoT), artificial intelligence (AI), and deep learning (DL) have been merged with blockchain technology to improve the old means of transportation. As a result of blockchain technology, the automotive sector is able to provide intelligent transportation systems and services, such as remote car operation and automated insurance.

Blockchain and Big Data in the Smart Grid Energy sector operations and processes could be considerably improved by blockchain technology. IoT systems and digital apps may be developed more quickly with the help of blockchain and big data, paving the way for P2P power trading and shared services. Modern energy systems are undergoing substantial changes as a result of the growth of distributed energy sources and the widespread use of information and communication technology. Control and management issues in suspended power systems and microgrids can be solved using blockchain frameworks (Gantz and Reinsel, 2011). Power networks may be made more efficient, reliable, and environmentally friendly through the use of intelligent grid technologies. As part of an intelligent grid, smart metres installed at the end of the user's metering infrastructure (AMI) ensure two connections between users and the equipment they are using. Many traditional solutions rely on a single entity to distribute and store keys in this process.

CONCLUSIONS

In recent years, blockchain has attracted a lot of attention as an innovative ledger technology that may enable big data systems while preserving high security and network efficiency. In this article, we present an in-depth update on the application of the big data blockchain technology. We began by discussing the current state of the blockchain and big data, as well as the reasons for combining the two technologies. A detailed assessment into the use of blockchain technology in a number of essential data services, including big-data collection, storage, analysis and protection has been carried out by our team of researchers. As a follow-up, we looked at how the blockchain can be used in big data applications such as intelligent transportation systems, intelligent healthcare systems, and intelligent grids. The development of big data platforms and the implementation of blockchain-based applications have also been noted and researched. We were able to identify many major technological problems and offer some potential routes for more research in this promising field after conducting a thorough assessment of the extant blockchain big data services and applications research literature.

To summarise, a solid blockchain infrastructure for big data has various technological issues that must be addressed during integration and implementation. This is evident from the discourse. Bigdata can help with blockchain problems, and the other way around as well. Big data's vast data accumulations and data services may be handled and

safeguarded with ease via the blockchain. Decentralized, immutable ledgers using current technology maintain data integrity and big data analytics provide more insights for generating meaningful projections based on massive data accumulations.

REFERENCES

1. *Alam, T., 2021. Blockchain-based big data integrity service framework for IoT devices data processing in smart cities. Mindanao Journal of Science and Technology.*
2. *C. H. Liu, Q. Lin, and S. Wen, "Blockchain-enabled data collection and sharing for industrial IoT with deep reinforcement learning," IEEE Transactions on Industrial Informatics, vol. 15, no. 6, pp. 3516–3526, Jun. 2019.*
3. *G. Liu, H. Dong, Z. Yan, X. Zhou, and S. Shimizu, "B4SDC: A blockchain system for security data collection in MANETs," IEEE Transactions on Big Data, 2020, in press.*
4. *H. Hu, Y. Wen, T.-S. Chua, and X. Li, "Toward scalable systems for big data analytics: A technology tutorial," IEEE access, vol. 2, pp. 652–687, 2014.*
5. *J. Gantz and D. Reinsel, "Extracting value from chaos," IDC iview, vol. 1142, no. 2011, pp. 1–12, Jun. 2011.*
6. *J. Manyika, "Big data: The next frontier for innovation, competition, and productivity," 2011. [Online]. Available: http://www.mckinsey.com/Insights/MGI/Research/Technology and Innovation/Big data The next frontier for innovation*
7. *J. Wu, K. Ota, M. Dong, J. Li, and H. Wang, "Big data analysis-based security situational awareness for smart grid," IEEE Transactions on Big Data, vol. 4, no. 3, pp. 408–417, Sep. 2018.*
8. *S. Pouyanfar, Y. Yang, S.-C. Chen, M.-L. Shyu, and S. Iyengar, "Multimedia big data analytics: A survey," ACM Computing Surveys (CSUR), vol. 51, no. 1, pp. 1–34, 2018.*
9. *X. Xu, X. Zhang, H. Gao, Y. Xue, L. Qi, and W. Dou, "BeCome: Blockchain-enabled computation offloading for IoT in mobile edge computing," IEEE Transactions on Industrial Informatics, vol. 16, no. 6, pp. 4187–4195, Jun. 2020.*
10. *Z. Su and Q. Xu, "Security-aware resource allocation for mobile social big data: A matching-coalitional game solution," IEEE Transactions on Big Data, 2017, in press.*
11. *Chavan, Amrita B., and K. Rajeswari. "The design and developement of decentralized digilocker using blockchain." International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR) 9 (2019): 29-36*
12. *Miraz, Mahadi Hasan, et al. "Factors Affecting Consumers Intention to Use Blockchain-Based Services (BBS) in the Hotel Industry." International Journal of Mechanical and Production Engineering Research and Development (IJMPERD) 10.3 (2020): 8891-8902.*
13. *Chaudhari, Mayuri V., Vrushali Choudhari, And Krantish Pol. "Cryptography Using Arithmetic Coding." International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR). 4. 2, Apr 2014, 1-4 © TJPRC Pvt. Ltd.*
14. *Bangar, Ashwini, and Swapnil Shinde. "Study and comparison of cryptographic methods for cloud security." Int J Comput Sci Eng Inf Technol Res 4.2 (2014): 205-213.*